Applied Statistics Education and Research Collaboration (ASEARC) - Conference Papers

Faculty of Engineering and Information Sciences

2012

Bayes Estimation in Meta-analysis using a linear model theorem

Junaidi
University of Newcastle

Elizabeth Stojanovski
University of Newcastle

Darfiana Nur
University of Newcastle

Bayes Estimation in Meta-analysis using a linear model theorem

Abstract

A Hierarchical Bayesian meta-analysis model developed by Dumouchel is derived by implementing the General Bayesian Linear model (GBLM) theorem. The aim is to obtain the joint posterior distribution of all parameters in the model. Simulation study is conducted to confirm the estimation of all parameters of interest. Results show parameter estimates as close to the true values indicating parameter stability.

Keywords

Meta-analysis, Bayes estimation, Hierarchical model, General Bayesian Linear Model theorem

Publication Details

Junaidi, Elizabeth Stojanovski and Darfiana Nur, Bayes Estimation in Meta-analysis using a linear model theorem, Proceedings of the Fifth Annual ASEARC Conference - Looking to the future - Programme and Proceedings, 2-3 February 2012, University of Wollongong.

Bayes Estimation in Meta-analysis using a linear model theorem

Junaidi[1,2], E. Stojanovski[1], D. Nur[1]

[1]The University of Newcastle, Callaghan, NSW, 2308, AUSTRALIA
[2]The University of Tadulako, Kampus Bumi Tadulako Tondo, Palu, 94118, INDONESIA
Junaidi@uon.edu.au, Elizabeth.Stojanovski@newcastle.edu.au, Darfiana.Nur@newcastle.edu.au

Abstract

A Hierarchical Bayesian meta-analysis model developed by Dumouchel is derived by implementing the General Bayesian Linear model (GBLM) theorem. The aim is to obtain the joint posterior distribution of all parameters in the model. Simulation study is conducted to confirm the estimation of all parameters of interest. Results show parameter estimates as close to the true values indicating parameter stability.

Keywords: Meta-analysis, Bayes estimation, Hierarchical model, General Bayesian Linear Model theorem.

1. Introduction

The analytical derivation of posterior distributions in the Bayesian framework can be computationally intractable due to the integration of multiple functions [10]. The Gibbs sampling algorithm is an iterative approach commonly used to calculate the joint posterior distribution of model parameters [5]. This algorithm works by simulating from the joint posterior distribution of model parameters and relies on decomposition of the joint posterior distribution into conditional distributions for each model parameter.

Assuming independence between parameters simplifies the choice of priors as independent sets of parameters may be considered separately. The selection of initial values is another important aspect to consider to ensure a suitable rate of convergence [4].

The Gibbs sampling approach is considered an efficient method for calculating the joint posterior distribution if parameter dependence is not high. Although derivation of the conditional distributions can be relatively easy, it is not always possible to find an efficient way to sample from these distributions [10].

An alternative approach to obtain the posterior distribution of model parameters is by application of the General Bayesian Linear Model (GBLM) theorem [8]. Multiple levels of a model are able to be relatively easily incorporated using the theorem to allow calculation of posterior distributions for more complex models [9]. The resulting posterior distribution is in the form of matrices. The theorem is presented in section 2.

The model under investigation in the present study is the hierarchical Bayesian model (HBM). This model has been used to accommodate model heterogeneity in meta-analysis [3]. A study [1] applied the HBM to account for heterogeneity between-trials in the form of differences in study and patient characteristics to determine whether a class of drugs reduces all-cause mortality in elderly patients with coronary heart disease (CHD). The HBM was also applied in another study [2] to estimate the posterior distribution of a standardised incidence ratio (SIR) to accommodate studies that reported the SIRs differently.

The aim of the present study is to calculate the posterior distribution for a HBM using the GBLM theorem. A simulation study is then conducted to estimate parameters of interest and assess parameter stability.

2. Methods

Hierarchical Bayesian Model

The standard HBM, proposed by Dumouchel [3], provides the following distributional assumptions:

$$Y_i \sim N(\theta_i, \sigma^2_{Y_i}) \quad i = 1,2,\ldots,n \qquad (1)$$

$$\theta_i \sim N(\mu, \sigma^2_\theta) \qquad (2)$$
$$\mu \sim N(0, D \rightarrow \infty).$$

Level 1 (1) of the model represents study data, where Y_i denotes the observed study statistics with variance-covariance matrix, $\sigma^2_{Y_i}$ and n (i =1,2,.,n) denotes the number of studies. Level 2 (2) represents the study-specific parameters, denoted θ and associated variance-covariance matrix σ^2_θ .. The overall mean is denoted by μ. The prior distribution for μ is assumed to follow the normal distribution with mean, zero, and variance $D \rightarrow \infty$ indicating elements of D as tending to infinity.

General Bayesian Linear Model Theorem

The GBLM theorem, developed by Lindley (1972) [8], is used here to obtain the posterior distribution of model parameters. The theorem is presented here in general terms. It is directly applied to the HBM in Section 3. This theorem requires a model to be presented in the form of matrices consistent with (3) to (9).

Suppose, given $\hat{\theta}_1$,

$$y \sim N\left(A_1\hat{\theta}_1, C_1\right), \qquad (3)$$

given $\hat{\theta}_2$,

$$\hat{\theta}_1 \sim N\left(A_2\hat{\theta}_2, C_2\right), \qquad (4)$$

and given $\hat{\theta}_3$,

$$\hat{\theta}_2 \sim N\left(A_3\hat{\theta}_3, C_3\right), \qquad (5)$$

Then $\boxed{\hat{\theta}_1 \mid \{A_k\}, \{C_k\}, \hat{\theta}_3, y \sim N(Dd, D)}$ (6)

with

$$D^{-1} = A_1^T C_1^{-1} A_1 + \{C_2 + A_2 C_3 A_2^T\}^{-1}, \qquad (7)$$

$$d = A_1^T C_1^{-1} y + \{C_2 + A_2 C_3 A_2^T\}^{-1} A_2 A_3 \hat{\theta}_3, \qquad (8)$$

and $\hat{\theta}_k$ a vector of p_k elements, the dispersion matrices, C_k, are assumed non-singular, y is a vector of n elements and A_k are known positive-definite matrices (k=1,2,3).

The joint distribution of $\hat{\theta}_1$and $\hat{\theta}_2$ is presented in (4) and (5). The use of the first part of the lemma presented in Lindley [8] enables the marginal distribution of $\hat{\theta}_1$ to be written as:

$$\hat{\theta}_1 \sim N(A_2A_3\hat{\theta}_3, C_2 + A_2C_3A_2^T). \qquad (9)$$

With (9) as a prior on $\hat{\theta}_1$ and the likelihood presented in (3), the second part of the lemma presented in Lindley [8] shows the posterior distribution of $\hat{\theta}_1$ as (6).

3. Results

Posterior Derivation

The posterior distribution of the Dumouchel model is now derived by applying the GBLM.

The first level of Dumouchel's model (1) is rewritten in matrix form according to the first stage of the GBLM (3).

$$\underbrace{\begin{bmatrix}Y_1\\Y_2\\Y_3\\\vdots\\Y_n\end{bmatrix}_{(n\,x\,1)}}_{y} \sim N\left(\underbrace{\begin{bmatrix}1&0&0&\cdots&0\\0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&0&0&1\end{bmatrix}_{(n\,x\,n)}}_{A_1}\underbrace{\begin{bmatrix}\theta_1\\\theta_2\\\theta_3\\\vdots\\\theta_n\end{bmatrix}_{(nx1)}}_{\hat{\theta}_1}, \underbrace{\begin{bmatrix}\sigma^2_{Y_1}&0&0&\cdots\\0&\sigma^2_{Y_2}&0&\cdots\\\vdots&\vdots&\ddots&\cdots\\0&0&\cdots&\sigma^2_{Y_n}\end{bmatrix}_{(nxn)}}_{C_1}\right) \quad (10)$$

The second level of the HBM (2) is rewritten according to the second stage of the GBLM (4).

$$\underbrace{\begin{bmatrix}\theta_1\\\theta_2\\\theta_3\\\vdots\\\theta_n\end{bmatrix}_{(n\,x\,1)}}_{\hat{\theta}_1} \sim N\left(\underbrace{\begin{bmatrix}1\\1\\1\\\vdots\\1\end{bmatrix}_{(nx1)}}_{A_2}\underbrace{[\mu]_{(1x1)}}_{\hat{\theta}_2}, \underbrace{\begin{bmatrix}\sigma^2_\theta&0&0&\cdots\\0&\sigma^2_\theta&0&\cdots\\\vdots&\vdots&\ddots&\cdots\\0&0&\cdots&\sigma^2_\theta\end{bmatrix}_{(nxn)}}_{C_2}\right) \quad (11)$$

Similarly for the prior of μ by applying (5).

$$\underbrace{[\mu]_{(1x1)}}_{\hat{\theta}_2} \sim N\left(\underbrace{[1]_{(1x1)}}_{A_3}\underbrace{[0]_{(1x1)}}_{\hat{\theta}_3}, \underbrace{[\sigma^2_\mu]_{(1x1)}}_{C_3}\right)$$

A_1 indicates unit matrix, A_2 is column vector with all elements 1. C_1, C_2 are dispersion matrices where the diagonal elements denote the variances and off-diagonal elements are zero; $\hat{\theta}_3$ is a zero due to the mean of the prior knowledge being zero. Column vectors are y, $\hat{\theta}_1$ all of the same dimension with elements of the form depicted in (10), (11) above. These vectors are assumed to follow the multivariate normal distributions.

By calculating D^{-1} and d for this model, using equations (7) and (8), respectively, we obtain the following:

$$D^{-1} = \begin{bmatrix} \sigma_{Y_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{Y_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_{Y_n}^2 \end{bmatrix}_{(nxn)}^{-1} + \begin{bmatrix} \sigma_\theta^2 + \sigma_\mu^2 & \sigma_\mu^2 & \cdots & \sigma_\mu^2 \\ \sigma_\mu^2 & \sigma_\theta^2 + \sigma_\mu^2 & \cdots & \sigma_\mu^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\mu^2 & \sigma_\mu^2 & \cdots & \sigma_\theta^2 + \sigma_\mu^2 \end{bmatrix}_{(nxn)}^{-1}$$

$$D^{-1} = \frac{1}{\sigma_\theta^2} \begin{bmatrix} 1 + \frac{\sigma_\theta^2}{\sigma_{Y_1}^2} - \frac{\sigma_\mu^2}{\sigma_\theta^2 + n\sigma_\mu^2} & -\frac{\sigma_\mu^2}{\sigma_\theta^2 + n\sigma_\mu^2} & \cdots & -\frac{\sigma_\mu^2}{\sigma_\theta^2 + n\sigma_\mu^2} \\ -\frac{\sigma_\mu^2}{\sigma_\theta^2 + n\sigma_\mu^2} & 1 + \frac{\sigma_\theta^2}{\sigma_{Y_2}^2} - \frac{\sigma_\mu^2}{\sigma_\theta^2 + n\sigma_\mu^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\sigma_\mu^2}{\sigma_\theta^2 + n\sigma_\mu^2} & -\frac{\sigma_\mu^2}{\sigma_\theta^2 + n\sigma_\mu^2} & \cdots & 1 + \frac{\sigma_\theta^2}{\sigma_{Y_n}^2} - \frac{\sigma_\mu^2}{\sigma_\theta^2 + n\sigma_\mu^2} \end{bmatrix}_{(nxn)} \quad (13)$$

and

$$d = \begin{bmatrix} Y_1 \frac{1}{\sigma_{Y_1}^2} & Y_2 \frac{1}{\sigma_{Y_2}^2} & Y_3 \frac{1}{\sigma_{Y_3}^2} & \cdots & Y_n \frac{1}{\sigma_{Y_n}^2} \end{bmatrix}_{1xn}^T . \quad (14)$$

Following matrix properties, specifically if a square matrix, D, is invertible, so is D^{-1}, and $(D^{-1})^{-1} = D$.
Using equation (6) of the GBLM theorem, we now have

$$\theta_1,...,\theta_n \mid \sigma_{Y_i}^2, \sigma_\theta^2, Y_1,...,Y_n \sim N(Dd, D) \quad (16)$$

where D and d are defined as the above.

Simulation study

A simulation study was conducted by generating 1,000 random samples, according to Dumouchel's model (1), (2), from the multivariate normal distribution for each of 30 studies using the R program. By fixing the overall mean and variance-covariance matrix, according to results from a previous case study [6], the estimated mean effect was calculated and compared to the true effect to assess bias. Steps for the simulation study comprised: fixing the value of the overall mean (μ) and variance-covariance matrices σ_{Yi}^2 and σ_θ^2; generating θ_i based on μ equalling 2.56; and obtaining values of the observed statistics (Y_i) based on θ_i.

Estimation of Parameters

The overall mean (μ) is estimated by applying the GBLM derivation of Dumouchel's model (6)(13)(14), to the simulated data (Y_i).

The first step was to read the observed statistics (Y_i) and variance-covariance matrix$(\sigma_{Y_i}^2)$ from the simulation study. The second step was to generate the posterior distribution (θ_i), which is assumed to follow the multivariate normal distribution, using the GBLM theorem (16). The value of the overall mean (μ) is estimated as the mean of this posterior distribution.

A thousand iterations were generated to estimate parameter. The results show the estimated value of the overall mean (μ = 2.54) as close to the true value (2.56). The 95% credible interval and standard deviation are 2.34–2.73 and 0.1, respectively. The histogram and density plot (Figures A and B, respectively) show the overall mean (μ) is within the credible interval consistently and close to the target (true) value.

Figure A. Histogram of μ

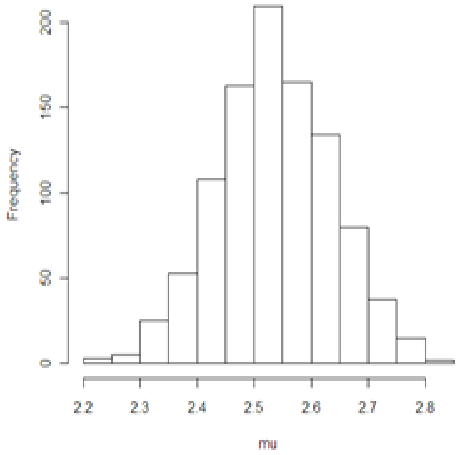
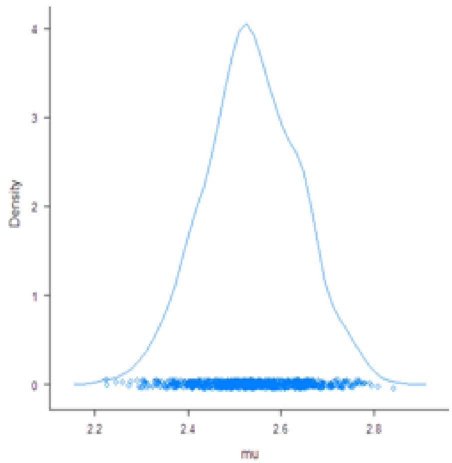


Figure B. Density of μ



4. Conclusion

The analytical form of a posterior distribution for a Bayesian model in meta-analysis has been presented. The posterior distribution is obtained using the GBLM theorem. The simulation study was conducted to estimate the overall mean. The results from the simulation study showed the overall estimated mean as close to the true value, confirming the estimator as consistent and stable.

5. Further Work

A case study should be further carried out to confirm the implementation of the model. Future investigation will also focus on applying the theorem to obtain posterior distributions of more complex model [7].

Acknowledgments

The first author in this research was supported by a Directorate General of Higher Education Indonesia (DIKTI) scholarship.

References

[1] Afilalo, J., Duque, G., et al., "Statins for secondary prevention in elderly patients - A hierarchical Bayesian meta-analysis." *Journal of the American College of Cardiology,* **51**(1): 37-45, 2008.

[2] Buja, A., Mastrangelo, G., Perissinotto, E., et al., "Cancer incidence among female flight attendants: A meta-analysis of published data." *Journal of Women's Health,* **15**: 98-105, 2006.

[3] Dumouchel, W., “Bayesian meta-analysis”, *Statistical Methodology in the Pharmaceutical Sciences*, 509-529, 1990.

[4] Fox, J-P., Glas, C. A. W., ”Bayesian Estimation of a multilevel IRT model using Gibbs sampling.” *Psychometrica*, **66**(2):271-288, 2001.

[5] Gilks, W. R., *Markov Chain Monte Carlo in Practice*, Chapman & Hall, London; Melbourne, 1996.

[6] Junaidi, Stojanovski, E. Nur, D., “Prior Sensitivity Analysis for a Hierarchical Model”, *Proceeding of the Fourth Annual ASEARCH Conference,* 64-67,2011.

[7] Jones, D. R., et. al., “Interspecies extrapolation in environmental exposure standard setting: A Bayesian synthesis approach.” *Regulatory Toxicology and Pharmacology*, **53**: 217-225, 2009.

[8] Lindley, D. V., Smith, A. F. M., ”Bayes Estimates for the Linear Model.” *Journal of the Royal Statistical Society,* **34**(1): 1-41, 2011.

[9] Welton, N. J., et. al., “Models for potentially biased evidence in meta-analysis using empirically based priors.” *Journal of the Royal Statistical Society*, **172**(1): 119-136, 2009.

[10] Tanner, M. A., *Tools for statistical inference: methods for the exploration of posterior distribution and likelihood function*, Springer, New York, 1996.